

2020

## Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* Spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series

Tatiana A. Rynearson  
*University of Rhode Island, rynearson@uri.edu*

Sarah A. Flickinger  
*University of Rhode Island*

Diana N. Fontaine  
*University of Rhode Island*

Follow this and additional works at: <https://digitalcommons.uri.edu/gsofacpubs>

---

### Citation/Publisher Attribution

Rynearson, T. A., Flickinger, S. A., & Fontaine, D. N. (2020). Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* Spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series. *Biology*, 9(1), 19. doi: 10.3390/biology9010019  
Available at: <http://dx.doi.org/10.3390/biology9010019>

This Article is brought to you for free and open access by the Graduate School of Oceanography at DigitalCommons@URI. It has been accepted for inclusion in Graduate School of Oceanography Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).



## Article

# Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* Spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series

Tatiana A. Ryneerson \*, Sarah A. Flickinger † and Diana N. Fontaine

Graduate School of Oceanography, University of Rhode Island, South Ferry Road, Narragansett, RI 02882, USA; sflickinger89@gmail.com (S.A.F.); fontained@uri.edu (D.N.F.)

\* Correspondence: ryneerson@uri.edu; Tel.: +1-(401)-874-6022

† Now at Catalog Technologies, 56 Roland Street, Suite 208, Boston, MA 02129, USA.

Received: 10 December 2019; Accepted: 7 January 2020; Published: 16 January 2020



**Abstract:** Diatoms generate nearly half of marine primary production and are comprised of a diverse array of species that are often morphologically cryptic or difficult to identify using light microscopy. Here, species composition and realized thermal niches of species in the diatom genus *Thalassiosira* were examined at the site of the Narragansett Bay (NBay) Long-Term Plankton Time Series using a combination of light microscopy (LM), high-throughput sequencing (HTS) of the 18S rDNA V4 region and historical records. *Thalassiosira* species were identified over 6 years using a combination of LM and DNA sequences. Sixteen *Thalassiosira* taxa were identified using HTS: nine were newly identified in NBay. Several newly identified species have small cell diameters and are difficult to identify using LM. However, they appeared frequently and thus may play a significant ecological role in NBay, particularly since their realized niches suggest they are eurythermal and able to tolerate the >25 °C temperature range of NBay. Four distinct species assemblages that grouped by season were best explained by surface water temperature. When compared to historical records, we found that the cold-water species *Thalassiosira nordenskiöldii* has decreased in persistence over time, suggesting that increasing surface water temperature has influenced the ecology of phytoplankton in NBay.

**Keywords:** community composition; diatom; diversity; high throughput sequencing; molecular ecology; phytoplankton; metabarcoding

## 1. Introduction

The combined activities of individual diatom species generate 40%–45% of oceanic primary production and influence the global cycling of silica and carbon [1–3]. Up to 200,000 diatom species are thought to exist [4,5], including species with different physiological and ecological characteristics including metabolic capability [6] and biogeographic range [5]. One of the challenges in accurately identifying the activities of individual species and their subsequent ecological and biogeochemical impacts is that many species are morphologically similar, or even identical, at the level of light microscopy [7,8].

Morphologically identical or "cryptic" species have been identified in the ecologically important diatom genus *Thalassiosira* [9,10]. Globally, *Thalassiosira* is one of the most abundant and diverse diatom genera [11] and includes species that are ecologically important components of phytoplankton communities in estuarine, coastal, and open ocean regions [12–17]. The identification of *Thalassiosira* species is known to be difficult, especially using light microscopy (LM) due to subtle differences in frustule morphology [18] and some *Thalassiosira* species are unidentifiable with LM methods [19].

While identification of *Thalassiosira* at the species level is possible using scanning electron microscopy (SEM) [13], this approach is generally not feasible for assessing large numbers of taxonomically complex field samples. Furthermore, small volume (e.g., 1 mL) counting methods typically used for LM identification of phytoplankton can underestimate species abundance by 20%–45% [20].

In many cases, community composition may be more accurately obtained through the use of molecular methods. For example, Hamsher et al. [10] were able to classify more *Thalassiosira* species from individual cells and colonies using molecular methods than LM, including rarely occurring species. Metabarcoding, or the high-throughput sequencing of PCR amplicons from an entire community, has been used to identify diatoms, including *Thalassiosira*, from complex communities [21–23]. At both the local and global scales, metabarcoding studies suggest that conventional LM methods have significantly underestimated diversity within the diatoms and they have revealed new and unidentified diatom ribotypes [21,24,25]. Generally, the focus of metabarcoding has been on either the entire phytoplankton community or on important functional groups, like diatoms, and most studies have reported taxonomy at the class, family and genus levels [22,25]. Few metabarcoding studies have focused on the ecological dynamics of individual species [21].

Here, we examined species richness and community composition in the diatom genus *Thalassiosira* using morphological count data and molecular data from the Narragansett Bay Long-Term Plankton Time Series (NBPTS) in Narragansett Bay (NBay), Rhode Island, USA [26,27]. The time series has collected weekly data on community composition since 1959. Historically, six *Thalassiosira* species were among the top 40 most abundant phytoplankton species in NBay [28]. *Thalassiosira nordenskiöldii*, the fifth most abundant phytoplankton species in NBay, was a characteristic member of the winter-spring bloom and *Thalassiosira rotula* was a characteristic fall blooming species [28]. With an understanding of the challenges involved in confidently identifying *Thalassiosira* to the species level, only a handful of taxa have regularly been recorded in the time series since 1959 (Table 1). This study used a metabarcoding approach combined with LM to investigate *Thalassiosira* community composition in NBay over a study period of 6 years and to address the following questions: (1) Which *Thalassiosira* species are present in NBay? (2) Are there temporal patterns in species composition and richness? (3) Do fluctuating environmental conditions drive observed temporal patterns? and (4) What are the realized thermal niches of species in NBay? To address these questions, we sequenced the 18S rDNA V4 region for 80 field samples collected over 6 years and compared this with species identification obtained using light microscopy. With this approach, *Thalassiosira* species not previously identified in NBay were identified and patterns of community composition were correlated with environmental conditions.

**Table 1.** Cell diameter and species range from Tomas [29], except where noted, for all *Thalassiosira* species identified with light microscopy (LM) (1959–1997 [27] and 1999–2014 [26]) and with high-throughput sequencing (HTS) (2008–2014). Species were either identified uniquely (x), identified as part of a group of indistinguishable species (\*), not detected because sequence not available in reference database (-) or not detected (blank). Numbers indicating the frequency of observation (no. of weeks observed) are included for rarely occurring species ( $\leq 10$  observations) in the light microscopy dataset.

Species	Cell Diameter ( $\mu\text{m}$ )	Geographic Range	LM 1959–1997	LM 1999–2014	HTS 2008–2014
<i>T. aestivalis</i>	14–56	Warm to temperate	7		*
<i>T. angulata</i>	12–39	Cold to temperate <sup>1</sup>			x
<i>T. anguste-lineata</i>	14–78	Cosmopolitan		1	x
<i>T. concaviuscula</i>	14–56 <sup>1</sup>	Neritic, cold to temperate <sup>1</sup>			x
<i>T. decipiens</i>	9–40	Cold to temperate	x		-
<i>T. eccentrica</i>	15–110	Cosmopolitan, except polar zones	2	1	x
<i>T. gravida</i>	17–62	Cold to temperate	x	*	*
<i>T. guillardii</i>	4–14 <sup>2</sup>	Marine and brackish <sup>2</sup>			x
<i>T. hyalina</i>	16–45	Cold to temperate	1		
<i>T. mala</i>	4–10	Warm to temperate			x
<i>T. minima</i>	5–15	Cosmopolitan, except polar zones			x
<i>T. nordenskiöldii</i>	10–50	Cold to temperate	x	x	*
<i>T. oceanica</i>	3–12	Mainly warm waters			x
<i>T. pacifica</i>	7–46	Cosmopolitan, except polar zones	9		*
<i>T. profunda</i>	1.8–5 <sup>3</sup>	Cosmopolitan <sup>2</sup>			x
<i>T. pseudonana</i>	2.3–5.5	Cosmopolitan	x		x
<i>T. punctigera</i>	40–186	Warm to temperate	8	x	x
<i>T. rotula</i>	8–55	Cosmopolitan	x	*	*
<i>T. subtilis</i>	15–32	Warm to temperate	5		-
<i>T. tenera</i>	10–29	Cosmopolitan <sup>4</sup>			x
<i>T. tumida</i>	21–137	Southern cold water region			x
<i>T. spp</i>			x	x	

<sup>1</sup> M Hoppenrath et al. [13]; <sup>2</sup> GR Hasle [19]; <sup>3</sup> J Belcher and E Swale [30]; <sup>4</sup> GR Hasle and G Fryxell [31].

## 2. Materials and Methods

### 2.1. Field Sampling, DNA Extraction and Sequencing

As part of the NBPTS, weekly surface water samples were collected between December 2008 and December 2014 from the west passage of NBay (41°34.2' N, 71°23.4' W), a partially mixed estuary in the northwest Atlantic. Water samples were filtered in triplicate onto 0.22 µm pore size, 25 mm diameter ExpressPlus filters (MilliporeSigma, Burlington, MA, USA) and stored at −80 °C for later DNA extraction. Filter volume was dependent on the in situ Secchi depth; 100 mL of water were filtered per 1 m of Secchi depth which ranged from 1 to 6 m.

Previously extracted DNA from 68 monthly surface water samples collected between December 2008 and December 2014 was used here [21]. An additional 12 monthly samples collected between January and December 2014 were selected based on monthly maximum *Thalassiosira* cell abundance and DNA was extracted from those frozen filters following Canesi and Ryneearson [21]. The final dataset contained 24 semi-monthly samples from 2014 and 80 field sampling dates overall (Table S1).

A 420 bp fragment within the variable V4 region of the 18S rDNA gene was amplified using primers D512 and D978rev [32] which amplify only diatoms and raphidophytes, enhancing our ability to amplify and detect rare diatom species. Primers were modified by the addition of Illumina specific adaptors: D512\_illumina: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTCCAGCTCCAATAGCG 3' and D978\_illumina: 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACGATGGTATCTAATC 3'. The following reagents were used in 10 µL PCR reactions; 1X Bio-x-Act Short Mix (Bioline USA Inc., Taunton, MA, USA), 0.5 µM each forward and reverse primer, and approximately 0.3–2.7 ng DNA template. Reactions were amplified with a multi-step thermocycler protocol, consisting of a 2 min denaturing step at 94 °C, followed by 20 cycles of 30 s each at 94, 49, and 72 °C, followed by 15 cycles of 30 s each at 94, 67, and 72 °C, followed by 10 min at 72 °C. Technical variation of this PCR reaction was measured through triplicate amplification and sequencing of one arbitrarily selected sample (30 December 2014).

PCR amplicons were cleaned with Ampure XP beads (Beckman Coulter, Inc., Brea, CA, USA), quantified with the Qubit High Sensitivity DNA Assay Kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA), amplified for an additional five cycles to add Nextera indices and adaptors (Illumina, Inc., San Diego, CA, USA) and cleaned again with Ampure XP beads. PCR products were pooled and quantified with the KAPA qPCR kit (Kapa Biosystems, Wilmington, MA, USA) prior to Illumina MiSeq sequencing with V2 chemistry (2 × 250 bp reads; Illumina, Inc., San Diego, CA, USA) at the University of Rhode Island Genomics and Sequencing Center.

### 2.2. Sequence Analysis

Paired end sequencing reads were first trimmed to remove low quality bases and Illumina adaptors using Trimmomatic PE [33], a 4 bp sliding window quality score of 20 and a minimum read length of 200 bp. Paired end sequences were merged using the BBmerge function from BBMap (Bushnell, Sourceforge.net/projects/bbmap/). Trimmed and merged sequences were filtered to a minimum length of 380 bp and a maximum expected error (based on phred quality scores) of 1 using USEARCH [34].

To identify taxa, filtered sequences were decomposed into minimum entropy decomposition (MED) nodes (equivalent to operational taxonomic units (OTUs)) [35], using a maximum entropy threshold of 0.0965 and a c value (maximum number of nucleotides with >0 entropy) of 4. Chimeric sequence nodes were identified and removed from results with UCHIME denovo [36] using an abundance skew value of 2.0. Taxonomy was assigned to remaining MED nodes using the BLAST method within the assign\_taxonomy.py script from QIIME [37] and a custom in-house reference database of 38 *Thalassiosira* sequences curated from GenBank using only sequences from taxonomically verified specimens or submitted by taxonomists. Discarded MED nodes included those not assigned a taxonomy within the *Thalassiosira* reference database and those that contained fewer than 1% of total sequence reads per sample to control for sequencing error and to obtain meaningful ecological information.

A maximum likelihood tree of the custom reference database was constructed using the curated database and two outgroup species (*Skeletonema marinoi* and *Skeletonema menzelii*). Sequences were aligned using MUSCLE [38]. The TrN + G model [39] was selected based on Bayesian Information Criterion using jModelTest2 [40,41]. The maximum likelihood tree was created with PhyML V3.0 and 100 bootstrap replicates in Geneious [42,43].

All statistical analyses of high-throughput sequencing (HTS) reads used only presence-absence data because relative abundance of sequence reads does not correspond to relative abundance of species [21]. Variation within triplicate PCR reactions was assessed using ANOVA. The probability of a species to occur on any given month was calculated by summing the presence of that species for each month over the time series and normalizing by sampling effort. A Bray–Curtis similarity matrix of species occurrence probabilities was used to generate a multidimensional scaling plot (MDS) in PRIMER v6.1.6 (PRIMER-E Ltd., Albany, Auckland, New Zealand).

### 2.3. Comparison of HTS Data With LM Counts and Environmental Analyses

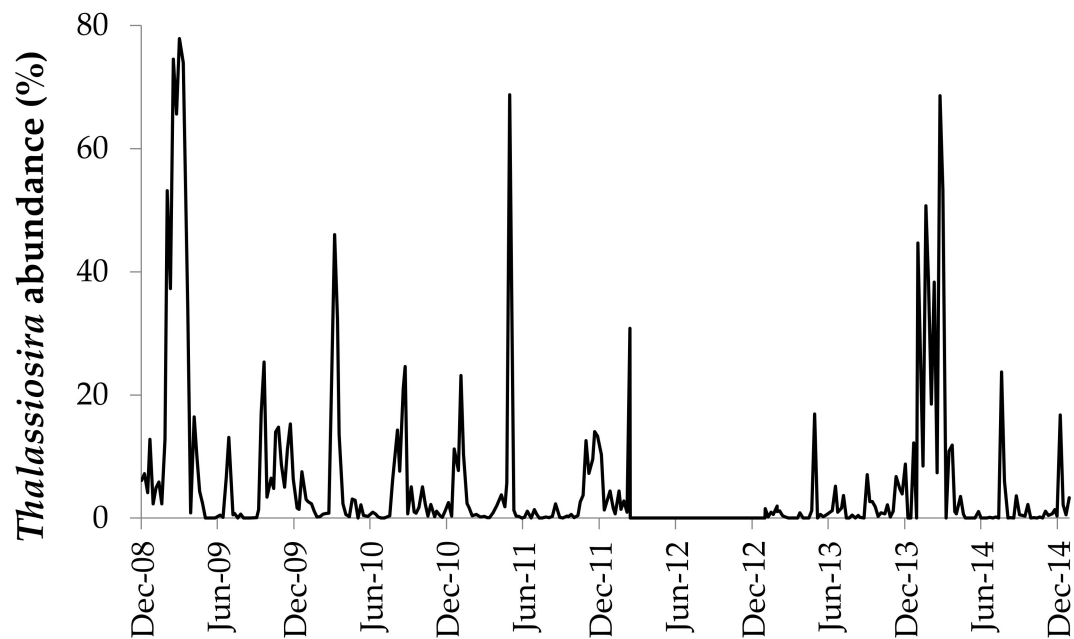
Light microscopy counts were obtained from the NBPTS [26] and the frequency and abundance of *Thalassiosira* species determined from December 2008 through December 2014 (Table S1). For seven dates across the 6-year dataset (23 May 2012, 25 June 2012, 2 July 2012, 1 August 2012, 5 September 2012, 5 October 2012, and 2 November 2012), light microscopy count data were unavailable. *Thalassiosira* species previously identified from NBay using LM were obtained from the modern (1999–present) [26] and historical time series (1959–1997) [27].

Species richness obtained using HTS was regressed against *Thalassiosira* abundance from the NBPTS using a least squares method and an ANOVA to test for significance of the slope. An ANOVA was used to test for significant differences in species richness with season (Winter: January–April, Spring: May–June, Summer: July–September, Fall: October–December). A multivariate trend correlation analysis (BIOENV) was used to correlate environmental factors with the Jaccard similarity index of community composition in PRIMER v6.1.6. Environmental data was obtained from the NBPTS (surface temperature, surface salinity, chlorophyll *a* [chl *a*], dissolved inorganic nitrogen [DIN], dissolved inorganic phosphorus [DIP], and silicate [Si]) [26]. For a small subset of samples (seven dates) for which chlorophyll *a* data was not available from the NBPTS, chl *a* data were obtained from the University of Rhode Island Marine Ecosystems Research Laboratory (MERL) Tank 98 bay sample dataset [44], which is collected weekly from the University of Rhode Island Graduate School of Oceanography dock (41.49°, −71.42°). This sample site is located 9.2 km south of Station II in the west passage of NBay and shares a similar physical environment with Station II. Additionally, temperature and salinity data for 15 September 2014 and nutrient data for 5 October 2014 were obtained from the MERL dataset [44]. Photosynthetically active radiation (PAR) data were obtained from the Narragansett Bay National Estuarine Research Reserve’s Narragansett Bay station (41.64°, −71.34°) [45]. Values for the seven environmental parameters were normalized in PRIMER v6.1.6 prior to analysis using BIOENV.

## 3. Results

### 3.1. Diversity and Occurrence of *Thalassiosira* Based on LM

Based on LM records in the long-term dataset, *Thalassiosira* species were present in 85% of the weekly surface water samples collected between December 2008 and December 2014 (Figure 1). During that same time period, *Thalassiosira* cell abundance ranged from 0 to 1,683,000 cells L<sup>−1</sup>, with a median abundance of 13,000 cells L<sup>−1</sup>. The largest blooms of *Thalassiosira* most often occurred in February and March and constituted up to 78% of the total microplankton in NBay. Four *Thalassiosira* species or groups of species were recorded in the long-term data set during this time period; two were identified to the species level (*Thalassiosira nordenskiöldii*, *Thalassiosira punctigera*) and two represented multiple species (*Thalassiosira rotula*/ *Thalassiosira gravida* and *Thalassiosira* spp.). The last two categories represent species that are difficult or impossible to distinguish at the resolution of light microscopy.

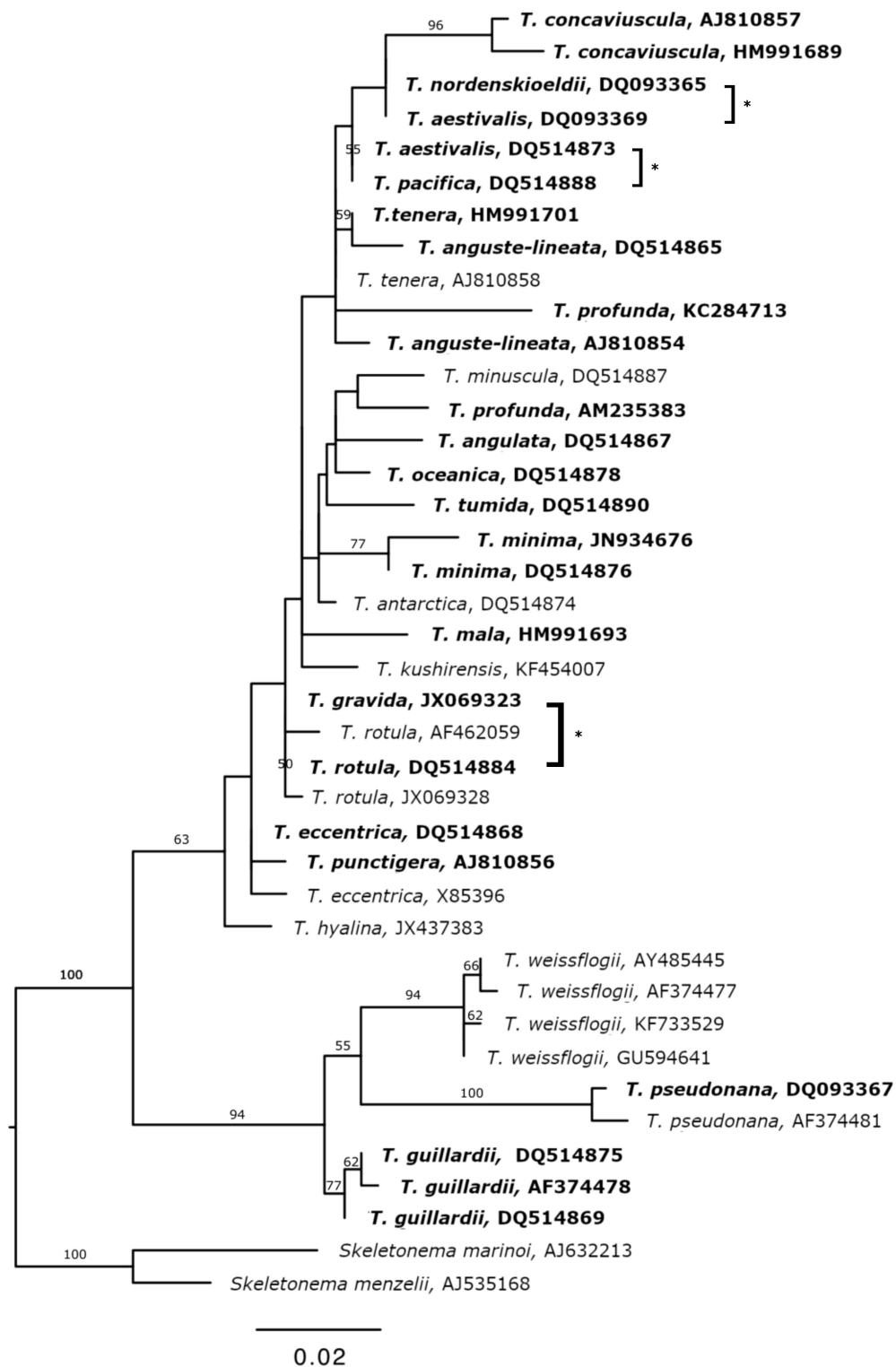


**Figure 1.** Percent abundance of *Thalassiosira* in weekly surface microplankton cell counts using light microscopy from the Narragansett Bay Long Term Plankton Time Series, December 2008–December 2014. Counts are not available from March–December 2012.

### 3.2. Reference Database

The 420 bp 18S rDNA V4 region did not clearly resolve all *Thalassiosira* species in the custom database (Figure 2). Five species were represented by multiple sequences with ambiguous placement on the tree, indicating that there are either species identification errors in GenBank or there is intraspecific variation at the V4 region for *Thalassiosira anguste-lineata*, *Thalassiosira profunda*, *Thalassiosira tenera*, *Thalassiosira eccentrica*, and *Thalassiosira aestivalis*. Three pairs of species could not be resolved at the 18S V4 locus because their sequences were not unique: *T. nordenskiöldii*/*T. aestivalis*, *Thalassiosira pacifica*/*T. aestivalis*, and *T. rotula*/*T. gravida*.





**Figure 2.** A maximum likelihood tree of *Thalassiosira* 18S V4 sequences included in the custom reference database. Bootstrap values (>50) are given at the nodes. Scale bar indicates genetic distance in substitutions per site. Bold labels denote sequences identified in the field samples analyzed here. Brackets with asterisks denote different species that are indistinguishable at the V4 region.



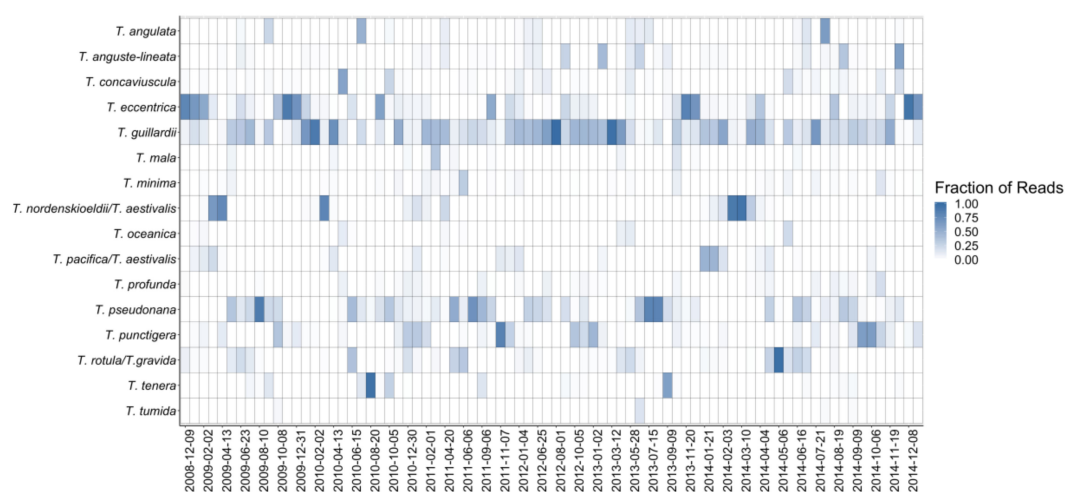
### 3.3. Sequencing Results

A total of  $13.6 \times 10^6$  read pairs from 82 samples (80 field samples and two amplification replicates) were sequenced (National Center for Biotechnology Information accession SRP078461). Reads per sample ranged from 58,448 to 281,234, with an average of 168,303 sequences per sample. An average of 13,016 sequences per sample (7.6%) survived the filtering and merging steps, with a range of 198 to 33,448. As the primers amplified all diatoms and raphidophytes, MED nodes that did not BLAST to our *Thalassiosira* reference database were discarded and not used in downstream analyses. Composition did not vary significantly among the three amplification triplicates (30 December 2014 field sample) (ANOVA,  $p = 0.74$ ). Among replicates, there was no difference in the rank order of species (Table S2). *Thalassiosira minima* occurred in low abundance (1.1%) in only one of the triplicates.

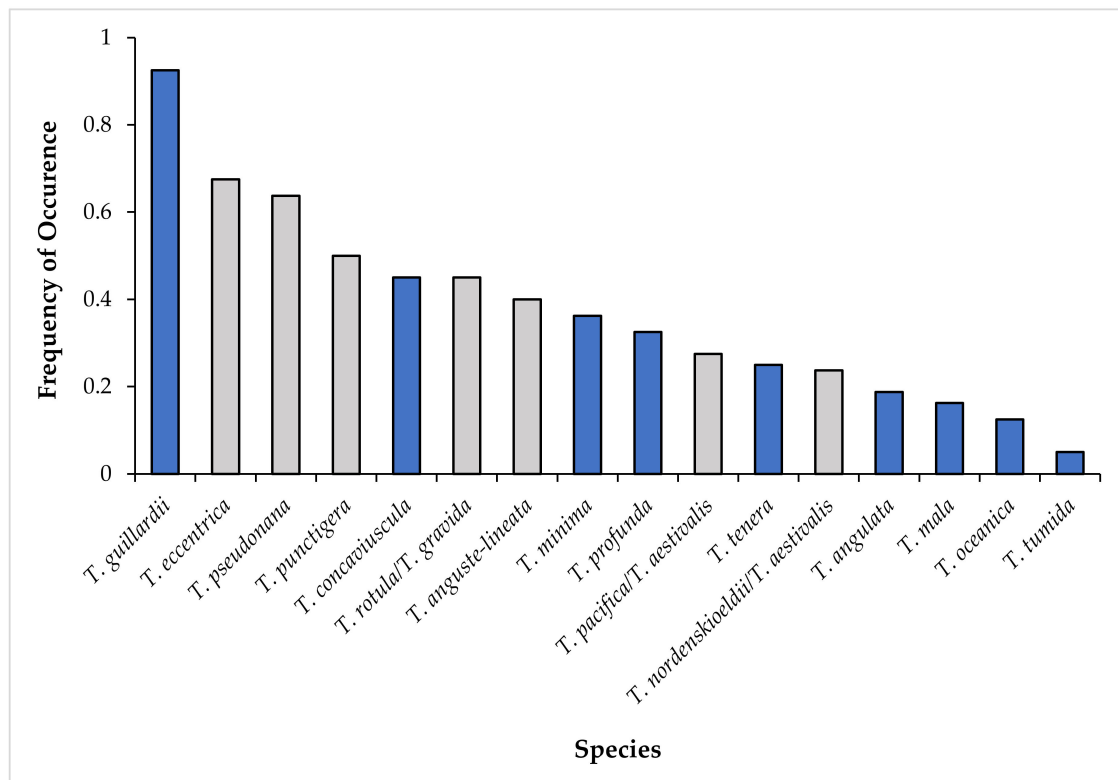
All 80 field samples contained positive read matches to the genus *Thalassiosira* even when LM count data indicated that no *Thalassiosira* species were present. Sixteen *Thalassiosira* taxa were identified in the HTS data including nine that had not been identified using LM (Table 1). Six of those taxa have cell diameters under 15  $\mu\text{m}$ , making them difficult to identify using LM at the 100 $\times$  magnification used to obtain counts as part of the time series.

### 3.4. Seasonality of *Thalassiosira* Community Composition

The 16 *Thalassiosira* species observed in the HTS dataset had different patterns (Figure 3) and frequencies of occurrence (Figure 4). The majority of species did not occur frequently; only three species appeared in more than 50% of field samples (Figure 4). *Thalassiosira guillardii* was the most frequent, appearing in 93% of field samples. *T. eccentrica* was the second most frequent (68%). *Thalassiosira pseudonana* was present in 64% of samples and all other species appeared less frequently.

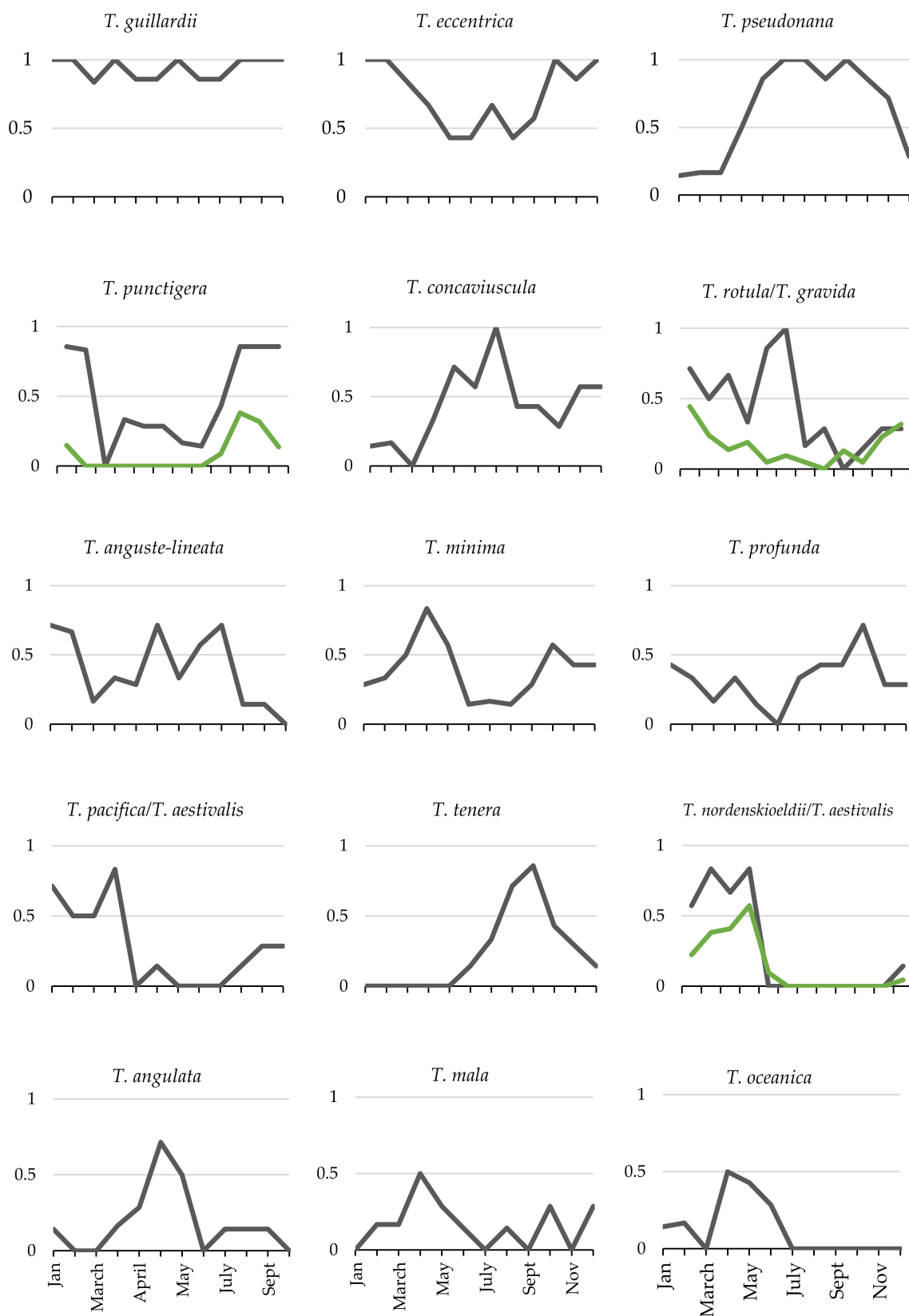


**Figure 3.** Heatmap showing relative abundance of sequence reads representing 16 *Thalassiosira* species identified using the metabarcoding approach. Samples were collected monthly between 2008 and 2013 and semi-monthly for 2014. Sample dates are represented in columns.



**Figure 4.** The frequency of occurrence of each species across all 80 sequenced field samples. Blue bars indicate newly identified species in Narragansett Bay.

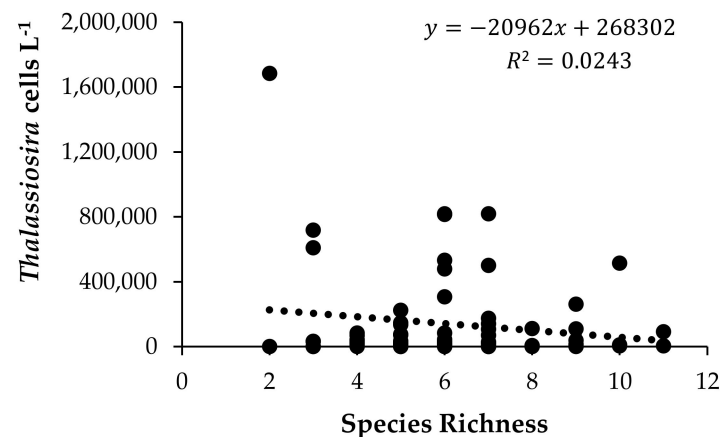
A variety of occurrence patterns was observed (Figure 5). The two most frequently occurring species, *T. guillardii* and *T. eccentrica*, had a high probability of occurrence throughout the year. In contrast, the third most frequently occurring species, *T. pseudonana*, had a strong seasonal signal with at least 85% occurrence during the spring and summer months (May–September). The remaining species generally had a seasonal signal of occurrence. One set of species had elevated probabilities of occurrence during the late fall and winter (*T. nordenskiöldii/T. aestivalis*, *T. pacifica/aestivalis*, and *Thalassiosira punctigera*). Another set of species occurred most frequently in late spring–summer (*T. rotula/T. gravida*, *Thalassiosira concaviuscula*, *Thalassiosira angulata*, and *Thalassiosira oceanica*). In mid-summer and fall, *T. tenera* and *Thalassiosira tumida* occurred more frequently whereas *Thalassiosira minima*, *Thalassiosira mala*, and *T. profunda* occurred more frequently in spring and fall.



**Figure 5.** Monthly probability of occurrence of each species in the sequenced field samples listed in order of abundance (cf. Figure 4). Green lines represent probability of occurrence data for species identified using LM. The least frequently occurring species, *Thalassiosira tumida*, is not shown.

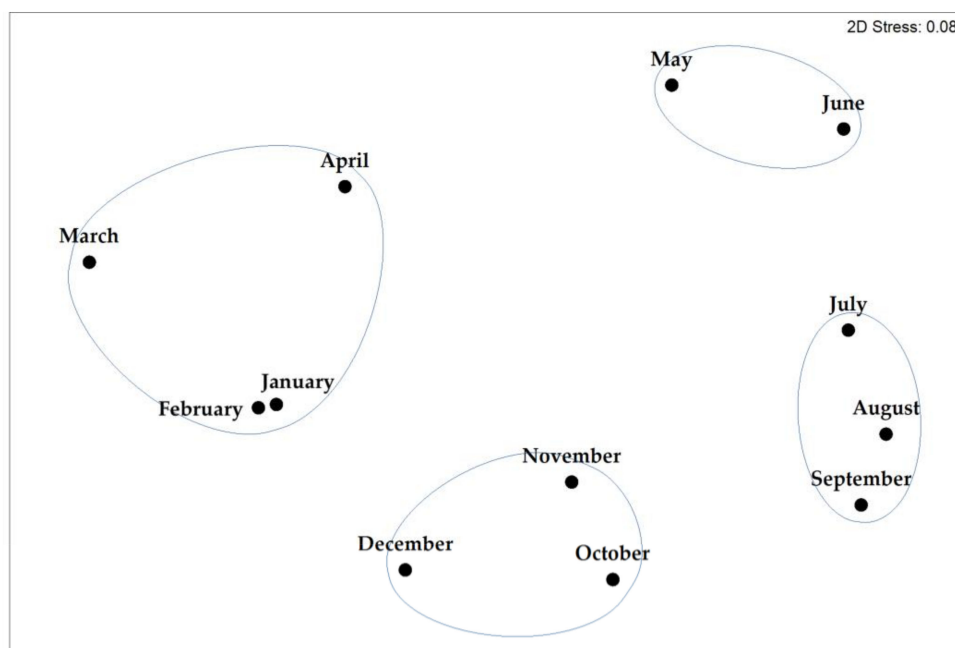
Notably, two of the species identified using both HTS and LM followed similar seasonal patterns (*T. punctigera*, *T. nordenskiöldii*, Figure 5). High-throughput sequencing and LM for *T. rotula/gravida* also followed roughly similar seasonal patterns with the exception that HTS indicated the highest probability of occurrence in June, a time when LM data suggested low levels of occurrence.

Species richness varied from 2 to 11 and some species did not co-occur. For example, the winter *T. nordenskiöldii*/*T. aestivalis* group never co-occurred with the late summer *T. tenera* or *T. tumida*. There was no significant relationship between species richness and the abundance of *Thalassiosira* obtained from the NBPTS ( $p = 0.18$ ) (Figure 6).



**Figure 6.** Species richness of *Thalassiosira* identified using HTS compared to cell counts obtained using LM for 80 field samples between 2008–2014.

To determine if there were similarities among species assemblages collected at different times of year, a Bray–Curtis similarity matrix of monthly species occurrence probabilities was calculated. A multidimensional scaling plot of the Bray–Curtis similarity matrix revealed four distinct seasonal assemblages: winter (January, February, March, April), spring (May, June), summer (July, August, September) and fall (October, November, December) (Figure 7). There was no significant difference in species richness among seasons ( $p = 0.63$ ) suggesting that while species composition shifted throughout the year, richness did not.



**Figure 7.** Multidimensional scaling plot (MDS) of Bray–Curtis similarity, based on the monthly probability of occurrence of each *Thalassiosira* species in the HTS dataset. Four seasonal groups are apparent at the 70% similarity level: winter (January–April), spring (May–June), summer (July–September), and fall (October–December).

### 3.5. Environmental Factors Associated With Species Occurrence

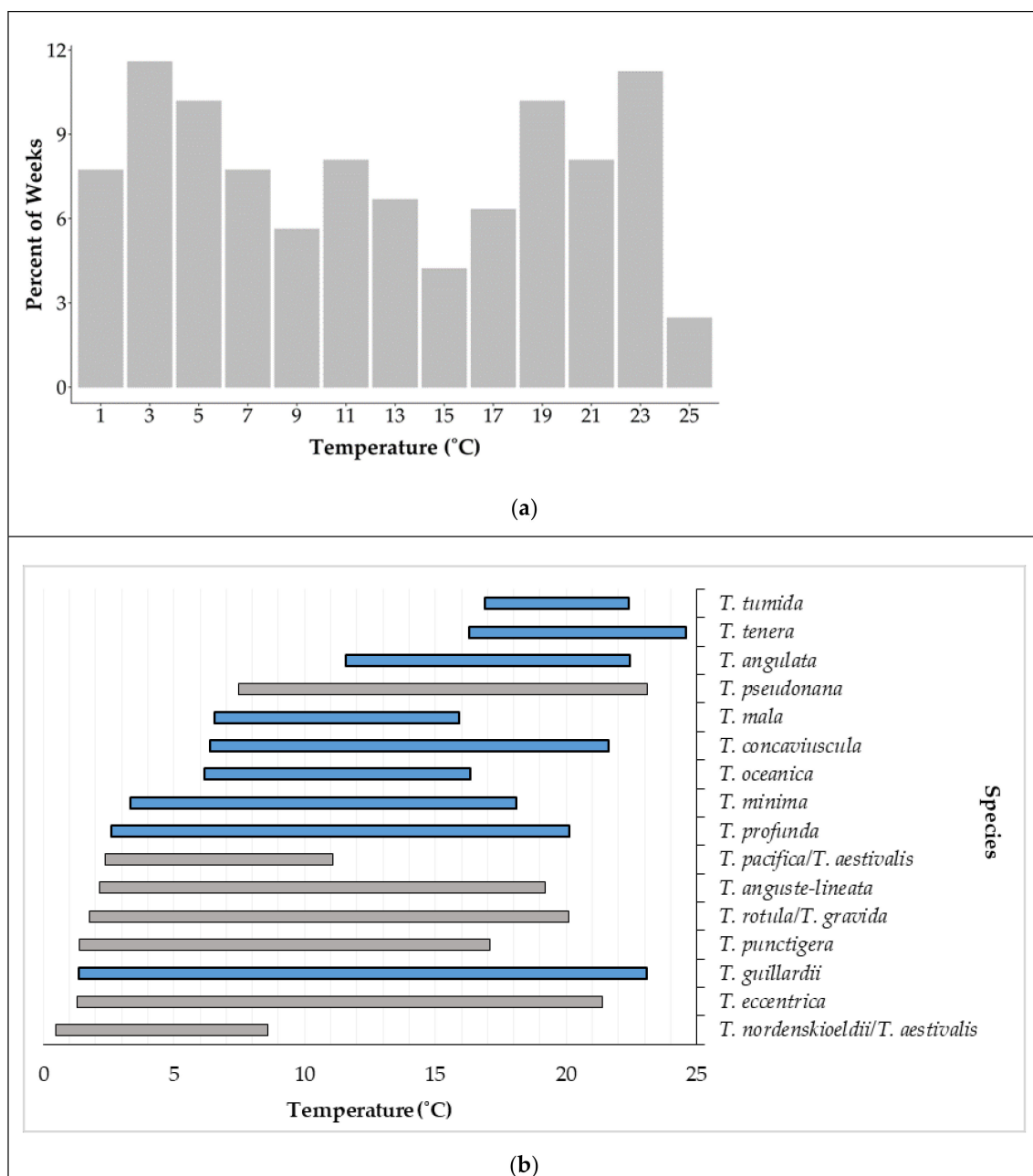
Environmental factors in NBay varied widely from December 2008 to December 2014 (Table 2, Figure S1). Among the variables examined, surface temperature explained 31.2% of the variation in *Thalassiosira* community composition (Table 3). The inclusion of additional variables decreased the correlation. Surface temperature in NBay from 2008 to 2014 showed a bimodal distribution with peaks between 2–4 and 22–24 °C which each occurred about 12% of the time (Figure 8a). The realized thermal niches of most species were broad, although a few were narrow. For example, the *T. nordenskiöldii*/*T. aestivalis* group only occurred between 0.5 and 8.5 °C (Figure 8b). *Thalassiosira mala* and *T. oceanica* were observed only at moderate temperatures (6.6–15.9 °C and 6.2–16.3 °C, respectively), while *T. tenera* and *T. tumida* were only observed at the warmest temperatures observed in NBay (16.3–24.6 °C).

**Table 2.** Range and median of environmental variables across the 80 field samples (2008–2014).

Environmental Variable	Range	Median
Sea surface temperature (°C)	0.5–24.6	13.9
Sea surface salinity	18.60–32.39	29.93
Average daily PAR (mmol m <sup>−2</sup> )	16.21–572.05	247.87
Dissolved inorganic nitrogen (μM)	0.14–15.80	2.61
Dissolved inorganic phosphorus (μM)	0.04–1.87	0.72
Silicate (μM)	0.05–38.12	11.46
Chlorophyll a (μg L <sup>−1</sup> )	0.23–29.80	5.14

**Table 3.** Multivariate correlation of *Thalassiosira* species composition with environmental factors.

Correlation	Environmental Variables
0.312	Surface Temperature
0.309	Surface Temperature, DIP
0.298	Surface Temperature, Si
0.296	Surface Temperature, Surface Salinity



**Figure 8.** (a) Frequency distribution of surface water temperatures, in 2 °C bins, measured at the Narragansett Bay Long Term Plankton Time Series site, December 2008–December 2014. Value on *x*-axis represents the middle of each bin (i.e., 1 °C for the >0–≤2 °C bin). (b) Observed temperature ranges for the *Thalassiosira* species recovered from the HTS data. Blue bars indicate newly identified species in Narragansett Bay.

#### 4. Discussion

##### 4.1. Utility of the 18S V4 Region for Species Identification in the Genus *Thalassiosira*

As few metabarcoding studies have focused on individual taxa and their ecological dynamics, we evaluated the utility of focused metabarcoding analysis of a single genus. By using diatom-specific primers that amplified the 18S rDNA V4 region [32], we identified *Thalassiosira* species more frequently using HTS than LM. *Thalassiosira* species were recovered from all 80 HTS samples but only from 70 samples using LM counts of 1 mL volumes. Notably, the metabarcoding data were obtained from

amplified DNA extracted from 100–300 mL filtered seawater, thus capturing species present at far less than 1 cell mL<sup>−1</sup>.

The 18S V4 region amplified here does not resolve all *Thalassiosira* species, since some have identical sequences at this locus (e.g., *T. nordenskioeldii* and *T. aestivalis*, *T. pacifica* and *T. aestivalis*, *T. rotula* and *T. gravida*). In total, our curated reference database differentiated 33 of 38 *Thalassiosira* species. Additionally, for some species in the genus, the 18S sequences are not present in public databases. For example, *Thalassiosira decipiens* is recorded in the NBPTS [27] but could not be detected in the sequence data because the 18S V4 sequence is not available in GenBank. This limitation is present for all barcoding loci used to date, including the 18S rDNA and highlights the need for continued efforts to obtain sequence data from taxonomically verified species.

Another limitation of the HTS metabarcoding approach is variation in 18S copy number among species. 18S gene copy number can vary from 1 to 12,000 copies cell<sup>−1</sup> in phytoplankton, and more specifically, can vary by an order of magnitude among *Thalassiosira* species [46]. For the large majority of *Thalassiosira* species, copy number per species is unknown. Although some have argued that abundances observed in LM and HTS efforts are comparable [25], experiments using mock communities revealed that relative abundance obtained using HTS differs significantly from in situ abundance [21,47]. Here, all statistical analyses and inferences were conducted using only presence-absence data to avoid the pitfalls associated with variable copy number.

#### 4.2. Species Diversity of *Thalassiosira* in Narragansett Bay

The diversity of *Thalassiosira* species in NBay is greater than previously realized. Of the 16 species identified using HTS, nine had not been historically documented in the NBPTS using LM. A combination of both HTS and LM data from 1959 through 2014 yielded a total of 20 *Thalassiosira* species or species groups (e.g., *T. rotula/gravida*). This is comparable to *Thalassiosira* species diversity found in other regions. For example, taxonomic studies using different combinations of LM, scanning electron microscopy, and DNA sequencing identified 21 *Thalassiosira* species in the North Sea [13] and 18 *Thalassiosira* species in a Scottish sea loch [16]. This suggests that even with an incomplete reference database, it is possible that a majority of *Thalassiosira* species in NBay were identified using the 18S V4 region.

Several of the newly recorded species in NBay are small (<15 µm in diameter) (*T. guillardii*, *T. mala*, *T. minima*, *T. oceanica*, *T. profunda*, *T. pseudonana*; Table 1) which may explain why they were not identified using LM at 100×, the magnification used to obtain species composition and abundance for the NBPTS. In fact, at least one of these small species, *T. profunda*, is not identifiable with LM [19]. The presence of these small species in NBay is consistent with their known habitats, which includes estuaries [16,30,48].

Interestingly, the second most abundant *Thalassiosira* species recorded in the historic dataset (and sixth most abundant phytoplankton species) [28] was characterized as a small (≈10 µm in diameter), unidentified *Thalassiosira* species that appeared in NBay beginning in 1967 and became both frequent and abundant. The most frequent species in the HTS data was the small *Thalassiosira guillardii* (4–14 µm in diameter). It was present during all months of the year and its probability of occurrence in any given month ranged from 83% (March) to 100% (multiple months). In other temperate estuaries, this species can occur frequently and in high cell numbers [30,48]. Other small diameter species in the HTS dataset, *T. minima* and *T. profunda*, appeared in >33% of all samples and have been observed previously from coastal habitats [30,49]. Given the small size of these species and the difficulty in distinguishing among them using LM, it is possible that a combination of species was present in the historic dataset and contributed to the highly abundant *Thalassiosira* spp. category. Together, they represent a set of small but potentially ecologically important diatoms in NBay, particularly given recent work highlighting the important role of nanoplanktonic diatoms in spring blooms and carbon export across the global ocean [50].



One of the smallest species identified in the HTS dataset is *T. pseudonana* (2.3–5.5  $\mu\text{m}$  diameter) [29]. Although it was not identified in NBPTS LM counts during this study period, it was the sixth most abundant *Thalassiosira* species recorded in the time series from 1959 to 1980 [28] and was present primarily in June and July [27]. In the HTS data, *T. pseudonana* was the third most frequently occurring species with highest occurrences between May and November (100% in June, July, and September), suggesting that this is primarily a warm water species in NBay, although its realized thermal niche spanned from 7.5 to 23  $^{\circ}\text{C}$ . *Thalassiosira pseudonana* can be a dominant spring taxon in Chesapeake Bay [51] and forms blooms in July and August in a brackish estuary off the Bay of Biscay [52]. This species is an important model organism with a fully sequenced genome [53]. Our results suggest that this species could be a more frequent and important contributor to estuarine phytoplankton communities than previously recognized.

A newly recorded, but larger, species in the dataset is *T. tumida*. Notably, this species has only been recorded from the Southern Ocean [29]. While this could suggest that the range of this species is greater than previously thought, samples with *T. tumida* were collected in the fall and summer when water temperatures were 16.9–22.4  $^{\circ}\text{C}$ . It is unlikely that the cause was either sequencing error since this taxon comprised 3%–15% of the sequence reads when it was present or misidentification in Genbank since multiple *T. tumida* accessions align and there are no other named species with a V4 region matching the *T. tumida* sequence exactly. A more likely scenario is that another, as yet unidentified *Thalassiosira* species shares the same 18S V4 sequence as *T. tumida* [9].

#### 4.3. Seasonal Patterns of Species Composition

Seasonal patterns were apparent both for individual species and whole communities of *Thalassiosira*. Multidimensional scaling analysis of the HTS data revealed four assemblages that corresponded with season. Of the environmental parameters that were analyzed, temperature was most highly correlated with variations in community composition. This result is consistent with other studies in NBay that have identified temperature as a key correlate of species composition including within the genus *Skeletonema* [21] and across all phytoplankton from 1959 to 1980 [28]. Narragansett Bay is a temperate estuary sited between the cold Boreal waters of the Gulf of Maine to the North and warmer waters of the mid-Atlantic Bight to the South and it experiences great excursions in temperature annually, from 0.5 to 24.6  $^{\circ}\text{C}$  during this study period.

Realized thermal niches of phytoplankton can be difficult to obtain given that many species cannot be detected using LM if cells are small and/or rare. The 6-year time frame of this study, broad temperature range in NBay, and metabarcoding approach allowed us to examine realized thermal niches in *Thalassiosira*, particularly for those species that were newly identified in the bay. The small diameter, newly identified *T. guillardii* could be characterized as eurythermal because it was present in all samples across all water temperatures. Alternately, the broad temperature range may be a sign that the *T. guillardii* V4 region represents a closely-related species complex, similar to that found in the diatom genus *Skeletonema*, which is comprised of many species with identical V4 sequences and in NBay, could appear eurythermal using the V4 [21]. Other species had realized thermal niches that were more restricted. At the coldest temperatures ( $\leq 1^{\circ}\text{C}$ ), only *T. nordenskiöldii* (see below) was present in the bay and at 6  $^{\circ}\text{C}$ , over half the *Thalassiosira* species observed in NBay were present. The warmest water temperatures had a similar pattern where only the newly identified *T. tenera* was present at temperatures above 23  $^{\circ}\text{C}$  but at water temperatures of 21  $^{\circ}\text{C}$ , about half of the species were present. Interestingly, surface waters were  $\leq 6^{\circ}\text{C}$  30% of the time and  $\geq 20^{\circ}\text{C}$  21% of the time, providing a substantial time frame for the proliferation of species that can survive at those temperatures. The range of realized thermal niches in NBay parallels observations by Thomas et al. [54] who found an exceptionally broad range of temperatures for optimal growth ( $\approx 6$ –30  $^{\circ}\text{C}$ ) in phytoplankton species collected from latitudes similar to NBay (41 $^{\circ}$  N).

*Thalassiosira nordenskiöldii* is a characteristic winter species in NBay [28]. Although *T. nordenskiöldii* could not be distinguished from *T. aestivalis* at the V4 region, the observed temperature range for the *T.*

*nordenskiöldii*/*T. aestivalis* group was 0.5–8.5 °C, which is within the temperature range (0.5–15 °C) identified for *T. nordenskiöldii* in NBay [28] and the laboratory [55]. In contrast, *T. aestivalis* has not generally been identified in the NBPTS aside from a handful of observations (seven times from 1959–1997) and is recognized as a temperate to warm water species [29], suggesting that the taxon identified here is likely *T. nordenskiöldii*. In an analysis of the NBPTS samples collected from 1959 to 1980, *Thalassiosira nordenskiöldii* was identified as the most abundant *Thalassiosira* species in NBay and the fifth most abundant diatom species overall [28]. In the 2008–2014 samples analyzed here, *T. nordenskiöldii* was not present at all from 2012 to 2013 in both the HTS and LM datasets. This species was present in 32% of the LM counts between 1959 and 1980 [28] but only in 15% of the LM counts during this study period. Given that the 2008–2014 time period of this study would likely capture one of the  $\approx 5$  year cycles in abundance that *T. nordenskiöldii* undergoes in NBay [28], the recent LM and HTS datasets suggest a decrease in persistence and possible decrease in ecological importance of this cold water species in NBay over the last 55 years. This is consistent with the nearly 2 °C warming of wintertime waters in NBay over the same time frame [56].

Other environmental factors that could significantly influence the composition and seasonality of *Thalassiosira* species in NBay include large-scale climatic processes, grazing by zooplankton, and spore formation. Large-scale processes such as the meandering of the Gulf Stream and the North Atlantic Oscillation are known to impact the size and timing of NBay summer *Skeletonema* blooms [57] and may also affect *Thalassiosira* species. Zooplankton grazing also has the potential to significantly alter community composition. Microzooplankton grazers consume an average of 96% of primary production in NBay [58] and the ctenophore *Mnemiopsis leidyi* can exert significant control on the plankton community in NBay [59]. Finally, community composition and seasonality can be significantly influenced by spore formation, which occurs in some *Thalassiosira* species [60] (reviewed in McQuoid and Hobson 1996). *Thalassiosira guillardii* is not known to form resting stages which may explain its persistence in the water column. In contrast, other frequently occurring species in NBay do form resting stages, including *T. nordenskiöldii*, *T. rotula*, *T. pseudonana*, and *T. minima* [60,61]. Together with large fluctuations in water temperature, resting stages may play an important role in the seasonality of species in NBay.

## 5. Conclusions

Using a combination of LM and metabarcoding approaches to analyze 6 years of data from the NBPTS, we found that *Thalassiosira* species diversity in NBay is greater than previously recognized. Much of this newly identified diversity can be attributed to several smaller sized species that are very frequent, have been overlooked using light microscopy, and may play an important ecological role. Overall, *Thalassiosira* species composition was most highly correlated with temperature. Analysis of realized thermal niches suggests that at least four types of species occur in NBay including cold, warm, and intermediate water species, as well as eurythermal species that can tolerate nearly a 25 °C temperature range. Given that temperatures in NBay are warming at a rate of nearly 0.2 °C per decade [56], there may be an increase in species that prefer warmer temperatures, such as *T. tenera*, and a continuation of the decrease we observed in persistence and possible ecological role of Arctic/Boreal species such as *T. nordenskiöldii*.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2079-7737/9/1/19/s1>, Table S1: List of field samples. Table S2: Species identified in sequencing triplicates. Figure S1: Monthly averages of environmental variables for the study period (2008–2014).

**Author Contributions:** Conceptualization, T.A.R., S.A.F.; methodology, S.A.F.; formal analysis, T.A.R., S.A.F., D.N.F.; investigation, S.A.F.; resources, T.A.R.; data curation, S.A.F., D.N.F.; writing—original draft preparation, T.A.R., S.A.F.; writing—review and editing, T.A.R., S.A.F., D.N.F.; visualization, S.A.F., D.N.F.; supervision, T.A.R.; project administration, T.A.R.; funding acquisition, T.A.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by National Science Foundation (NSF) awards 1638834 and 1655686 (to TAR), the University of Rhode Island, via their ongoing support for the Narragansett Bay Long-Term Plankton

Time Series and NSF award OIA-1655221 which partially supported DNF. This material is based upon work conducted at two Rhode Island NSF EPSCoR research facilities, the Genomics and Sequencing Center, University of Rhode Island and the Center for Computation and Visualization, Brown University, both supported in part by the National Science Foundation EPSCoR awards OIA-1004057 and OIA-1655221.

**Acknowledgments:** Special thanks to the Plankton Assistants for collecting samples and generating LM data, G. Chen for his assistance in submitting sequence data to NCBI, and O. Ahern for helpful suggestions regarding data analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Field, C.B.; Behrenfeld, M.J.; Randerson, J.T.; Falkowski, P. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **1998**, *281*, 237–240. [\[CrossRef\]](#)
- Tréguer, P.J.; De La Rocha, C.L. The world ocean silica cycle. *Annu. Rev. Mar. Sci.* **2013**, *5*, 477–501. [\[CrossRef\]](#)
- Rousseaux, C.; Gregg, W. Interannual variation in phytoplankton primary production at a global scale. *Remote Sens.* **2014**, *6*, 1–19. [\[CrossRef\]](#)
- Mann, D.G.; Droop, S.J.M. Biodiversity, biogeography and conservation of diatoms. *Hydrobiologia* **1996**, *336*, 19–32. [\[CrossRef\]](#)
- Mann, D.G.; Vanormelingen, P. An inordinate fondness? The number, distributions, and origins of diatom species. *J. Eukaryot. Microbiol.* **2013**, *60*, 414–420. [\[CrossRef\]](#)
- Alexander, H.; Jenkins, B.D.; Rynearson, T.A.; Dyhrman, S.T. Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E2182–E2190. [\[CrossRef\]](#)
- Amato, A.; Kooistra, W.H.C.F.; Levialdi Ghiron, J.H.; Mann, D.G.; Proschold, T.; Montresor, M. Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* **2007**, *158*, 193–207. [\[CrossRef\]](#)
- Beszteri, B.; Acs, E.; Medlin, L.K. Ribosomal DNA sequence variation among sympatric strains of the *Cyclotella meneghiniana* complex (Bacillariophyceae) reveals cryptic diversity. *Protist* **2005**, *156*, 317–333. [\[CrossRef\]](#) [\[PubMed\]](#)
- Whittaker, K.; Rignanes, D.; Olson, R.; Rynearson, T. Molecular subdivision of the marine diatom *Thalassiosira rotula* in relation to geographic distribution, genome size, and physiology. *BMC Evol. Biol.* **2012**, *12*, 209. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hamsher, S.; LeGresley, M.; Martin, J.; Saunders, G. A comparison of morphological and molecular-based surveys to estimate the species richness of *Chaetoceros* and *Thalassiosira* (Bacillariophyta), in the Bay of Fundy. *PLoS ONE* **2013**, *8*, e73521. [\[CrossRef\]](#) [\[PubMed\]](#)
- Leblanc, K.; Aristegui, J.; Armand, L.; Assmy, P.; Beker, B.; Bode, A.; Breton, E.; Cornet, V.; Gibson, J.; Gosselin, M.P.; et al. A global diatom database-abundance, biovolume and biomass in the world ocean. *Earth Syst. Sci. Data* **2012**, *4*, 149–165. [\[CrossRef\]](#)
- Pratt, D.M. The winter-spring diatom flowering in Narragansett Bay. *Limnol. Oceanogr.* **1965**, *10*, 173–184. [\[CrossRef\]](#)
- Hoppenrath, M.; Beszteri, B.; Drebes, G.; Halliger, H.; Van Beusekom, J.E.E.; Janisch, S.; Wiltshire, K.H. *Thalassiosira* species (Bacillariophyceae, *Thalassiosirales*) in the North Sea at Helgoland (German Bight) and Sylt (North Frisian Wadden Sea) – a first approach to assessing diversity. *Eur. J. Phycol.* **2007**, *42*, 271–288. [\[CrossRef\]](#)
- Yoshie, N.; Suzuki, K.; Kuwata, A.; Nishioka, J.; Saito, H. Temporal and spatial variations in photosynthetic physiology of diatoms during the spring bloom in the western subarctic Pacific. *Mar. Ecol. Prog. Ser.* **2010**, *399*, 39–52. [\[CrossRef\]](#)
- Muyllaert, K.; Sabbe, K. The diatom genus *Thalassiosira* (Bacillariophyta) in the estuaries of the Schelde (Belgium/The Netherlands) and the Elbe (Germany). *Bot. Mar.* **1996**, *36*, 103–115. [\[CrossRef\]](#)
- Harris, A.; Medlin, L.; Lewis, J.; Jones, K. *Thalassiosira* species (Bacillariophyceae) from a Scottish sea-loch. *Eur. J. Phycol.* **1995**, *30*, 117–131. [\[CrossRef\]](#)
- Aizawa, C.; Tanimoto, M.; Jordan, R. Living diatom assemblages from North Pacific and Bering Sea surface waters during summer 1999. *Deep Sea Res. II* **2005**, *52*, 2186–2205. [\[CrossRef\]](#)
- Kaczmarek, I.; Beaton, M.; Benoit, A.C.; Medlin, L.K. Molecular phylogeny of selected members of the order *Thalassiosirales* (Bacillariophyta) and evolution of the fultoportula. *J. Phycol.* **2006**, *42*, 121–138. [\[CrossRef\]](#)

19. Hasle, G.R. Some freshwater and brackish water species of the diatom genus *Thalassiosira* Cleve. *Phycologia* **1978**, *17*, 263–292. [CrossRef]
20. Rodriguez-Ramos, T.; Dornelas, M.; Maranon, E.; Cermeno, P. Conventional sampling methods severely underestimate phytoplankton species richness. *J. Plankton Res.* **2013**, *36*, 334–343. [CrossRef]
21. Canesi, K.; Rynearson, T.A. Temporal variation of *Skeletonema* community composition from a long-term time series in Narragansett Bay identified using high-throughput sequencing. *Mar. Ecol. Prog. Ser.* **2016**, *556*, 1–16. [CrossRef]
22. De Vargas, C.; Audic, S.; Henry, N.; Decelle, J.; Mahé, F.; Logares, R.; Lara, E.; Berney, C.; Le Bescot, N.; Probert, I.; et al. Eukaryotic plankton diversity in the sunlit ocean. *Science* **2015**, *348*. [CrossRef] [PubMed]
23. Endo, H.; Ogata, H.; Suzuki, K. Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean. *Sci. Rep. UK* **2018**, *8*, 10916. [CrossRef] [PubMed]
24. Piredda, R.; Claverie, J.-M.; Decelle, J.; de Vargas, C.; Dunthorn, M.; Edvardsen, B.; Eikrem, W.; Forster, D.; Kooistra, W.H.C.F.; Logares, R.; et al. Diatom diversity through HTS-metabarcoding in coastal European seas. *Sci. Rep.* **2018**, *8*, 18059. [CrossRef] [PubMed]
25. Malviya, S.; Scalco, E.; Audic, S.; Vincent, F.; Veluchamy, A.; Poulain, J.; Wincker, P.; Iudicone, D.; de Vargas, C.; Bittner, L.; et al. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1516–E1525. [CrossRef] [PubMed]
26. Narragansett Bay Long-Term Plankton Time Series. Available online: <https://web.uri.edu/gso/research/plankton/> (accessed on 29 May 2016).
27. Smayda, T.J.; The Bunker C Community. Narragansett Bay Plankton Time Series (1959–1997). Available online: <http://www.nabats.org/nabats-phytoplankton-data.html> (accessed on 1 November 2019).
28. Karentz, D.; Smayda, T.J. Temperature and seasonal occurrence patterns of 30 dominant phytoplankton species in Narragansett Bay over a 22-year period (1959–1980). *Mar. Ecol. Prog. Ser.* **1984**, *18*, 277–293. [CrossRef]
29. Tomas, C.R. *Identifying Marine Phytoplankton*; Academic Press: New York, NY, USA, 1997; p. 858.
30. Belcher, J.; Swale, E. Notes on some small *Thalassiosira* species (Bacillariophyceae) from the plankton of the lower Thames and other British Estuaries (identified by transmission electron microscopy). *Br. Phycol. J.* **1986**, *21*, 139–145. [CrossRef]
31. Hasle, G.R.; Fryxell, G. The genus *Thalassiosira*: Some species with a linear areola array. *Beih. Noca Hedwig.* **1977**, *54*, 15–66.
32. Zimmermann, J.; Jahn, R.; Gemeinholzer, B. Barcoding diatoms: Evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Org. Divers. Evol.* **2011**, *11*, 173–192. [CrossRef]
33. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]
34. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [CrossRef] [PubMed]
35. Eren, A.M.; Morrison, H.G.; Lescault, P.J.; Reveillaud, J.; Vineis, J.H.; Sogin, M.L. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J.* **2015**, *9*, 968–979. [CrossRef] [PubMed]
36. Edgar, R.C.; Haas, B.J.; Clemente, J.C.; Quince, C.; Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **2011**, *27*, 2194–2200. [CrossRef] [PubMed]
37. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Pena, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* **2010**, *7*, 335–336. [CrossRef] [PubMed]
38. Edgar, R.C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **2004**, *5*, 113. [CrossRef] [PubMed]
39. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. [CrossRef]
40. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]
41. Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **2003**, *52*, 696–704. [CrossRef]



42. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyl 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [\[CrossRef\]](#)
43. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [\[CrossRef\]](#)
44. University of Rhode Island Marine Ecosystems Research laboratory. Available online: <http://www.gso.uri.edu/merl/data.htm> (accessed on 29 May 2016).
45. National Estuarine Research Reserve Centralized Data Management Office. Available online: <http://cdmo.baruch.sc.edu/> (accessed on 29 May 2016).
46. Zhu, F.; Massana, R.; Not, F.; Marie, D.; Vault, D. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* **2005**, *52*, 79–92. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Mäki, A.; Salmi, P.; Mikkonen, A.; Kremp, A.; Tirola, M. Sample preservation, DNA or RNA extraction and data analysis for high-throughput phytoplankton community sequencing. *Front. Microbiol.* **2017**, *8*. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Maria Trigueros, J.; Orive, E. Seasonal variations of diatoms and dinoflagellates in a shallow, temperate estuary, with emphasis on neritic assemblages. *Hydrobiologia* **2001**, *444*, 119–133. [\[CrossRef\]](#)
49. Hasle, G.R. The biogeography of some marine planktonic diatoms. *Deep Sea Res.* **1976**, *23*, 319–338.
50. Leblanc, K.; Quéguiner, B.; Diaz, F.; Cornet, V.; Michel-Rodriguez, M.; Durrieu de Madron, X.; Bowler, C.; Malviya, S.; Thyssen, M.; Grégori, G.; et al. Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. *Nat. Commun.* **2018**, *9*, 953. [\[CrossRef\]](#)
51. Marshall, H.G.; Lacouture, R.V.; Buchanan, C.; Johnson, J.M. Phytoplankton assemblages associated with water quality and salinity regions in Chesapeake Bay, USA. *Estuar. Coast. Shelf Sci.* **2006**, *69*, 10–18. [\[CrossRef\]](#)
52. Hevia-Orube, J.; Orive, E.; David, H.; Díez, A.; Laza-Martínez, A.; Miguel, I.; Seoane, S. Molecular and morphological analyses of solitary forms of brackish Thalassiosiroid diatoms (Coscinodiscophyceae), with emphasis on their phenotypic plasticity. *Eur. J. Phycol.* **2016**, *51*, 11–30. [\[CrossRef\]](#)
53. Armbrust, E.V.; Berges, J.A.; Bowler, C.; Green, B.R.; Martinez, D.; Putnam, N.H.; Zhou, S.G.; Allen, A.E.; Apt, K.E.; Bechner, M.; et al. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science* **2004**, *306*, 79–86. [\[CrossRef\]](#)
54. Thomas, M.K.; Kremer, C.T.; Klausmeier, C.A.; Litchman, E. A global pattern of thermal adaptation in marine phytoplankton. *Science* **2012**, *338*, 1085–1088. [\[CrossRef\]](#)
55. Durbin, E.G. Studies on the autoecology of the marine diatom *Thalassiosira nordenskiöldii* Cleve. 1. The influence of daylength, light intensity, and temperature on growth. *J. Phycol.* **1974**, *10*, 220–225. [\[CrossRef\]](#)
56. Fulweiler, R.W.; Oczkowski, A.J.; Miller, K.M.; Oviatt, C.A.; Pilson, M.E.Q. Whole truths vs. half truths – And a search for clarity in long-term water temperature records. *Estuar. Coast. Shelf Sci.* **2015**, *157*, A1–A6. [\[CrossRef\]](#)
57. Borkman, D.G.; Smayda, T.J. Gulf Stream position and winter NAO as drivers of long-term variations in the bloom phenology of the diatom *Skeletonema costatum* “species-complex” in Narragansett Bay, RI, USA. *J. Plankton. Res.* **2009**, *31*, 1407–1425. [\[CrossRef\]](#)
58. Lawrence, C.M.; Menden-Deuer, S. Drivers of protistan grazing pressure: Seasonal signals of plankton community composition and environmental conditions. *Mar. Ecol. Prog. Ser.* **2012**, *459*, 39–52. [\[CrossRef\]](#)
59. Deason, E.; Smayda, T. Ctenophore-zooplankton-phytoplankton interactions in Narragansett Bay, Rhode Island, USA, during 1972–1977. *J. Plankton. Res.* **1981**, *4*, 203–217. [\[CrossRef\]](#)
60. McQuoid, M.R.; Hobson, L.A. Diatom resting stages. *J. Phycol.* **1996**, *32*, 889–902. [\[CrossRef\]](#)
61. McQuoid, M.R. Influence of salinity on seasonal germination of resting stages and composition of microplankton on the Swedish west coast. *Mar. Ecol. Prog. Ser.* **2005**, *289*, 151–163. [\[CrossRef\]](#)

